The Role of Object Shadows in Promoting 3D Visualization

GEOFFREY S. HUBONA[1]
Department of Information Systems
Virginia Commonwealth University, Richmond, VA 23284
PHILIP N. WHEELER
Century Computing, Inc., 8101 Sandy Spring Road, Laurel, MD 20707
GREGORY W. SHIRAH and MATTHEW BRANDT
Goddard Space Flight Center, Code 522, Greenbelt, MD 20771

More powerful contemporary computer hardware has enabled a wider variety of techniques to depict computer-generated objects in three dimensional (3D) space. The use of stereoscopic viewing and motion have been extensively studied. However, the effective use of computer-rendered object shadows to provide spatial information about the relative position and size of objects in virtual space has not. Subjects perform two tasks with 3D geometric patterns of objects presented on a computer screen: (1) positioning the object to complete a symmetrical geometric figure; and (2) resizing the object to match the size of other objects. Performance accuracy and speed are recorded under the following conditions: (1) objects casting shadows on and off; (2) one or two light sources; (3) stereoscopic and monoscopic viewing; and (4) different backgrounds for the cast shadows: flat plane (i.e. floor); 'stair-step' floor with no walls; and floor with walls (i.e. room). The use of object shadows as depth cues enhances positioning performance, but not resizing performance. Moreover, the object shadows are not as effective as stereoscopic viewing in facilitating performance. Furthermore, performance degrades with the stair-step shadow background, and when the number of light sources increases from one to two.

Categories and Subject Descriptors: H.1 [**Models and Principles**]: User/Machine Systems; H.5 [**Information Interfaces and Presentation**]: User Interfaces; I.3 [**Computer Graphics**]: Three-Dimensional Graphics and Realism

General Terms: Experimentation, Human Factors

Additional Key Words and Phrases: 3D user interfaces, shadows, depth perception, cue theory

---

1. INTRODUCTION

There has been a dramatic increase in the performance capabilities of personal computers and modern workstations in recent years. The increasing power of computing machinery has enabled the proliferation of three-dimensional (3D) user interface techniques for a variety of business, entertainment, and scientific applications. However, a key challenge facing 3D interface designers is to develop effective techniques to depict objects in 3D space on a physical medium that is inherently two-dimensional (2D): a flat computer screen. It is especially important to depict spatial relationships among objects in 3D space, particularly with respect to depth (e.g. the *z* dimension) so that users can locate and manipulate these objects. To this end, interface designers use various depth cues to make objects 'appear' 3D on a 2D screen. These techniques are based on psychological research about human perception and the cognitive processing of visual information.

---

[1] Primary author's current address is: 2215 Hanover Avenue, 1st Floor, Richmond, VA 23220

The visual cues that enable humans to perceive depth have been extensively documented. However, the use of these cues, and particularly, the combinations of these cues, to effectively convey depth information in computer-generated scenes remains an ongoing topic of research. Often depth cues are placed in two categories: (1) primary cues including binocular disparity, convergence and accommodation; and (2) secondary cues that include perspective and elevation, size, texture, shading and shadow, motion, reference frames, and others (see Kelsey [1993] for a complete discussion of these primary and secondary depth cues).

Binocular disparity is a primary, physiological characteristic that enables the stereoscopic viewing of objects within a limited distance. Stereo viewing facilitates spatial tasks by providing powerful depth information about the relative sizes and positions of objects in space. However, there are practical limitations with respect to the use of stereo vision-enabling hardware and software in computer applications. It is relatively expensive, requires the concomitant use of often cumbersome head-mounted apparatuses, and it is not supported by most existing software. For these reasons, it is useful to research visual techniques that may serve as practical alternatives to stereo viewing in conveying depth information.

The use of real time, moving object shadows to convey depth information about the corresponding objects in 3D space is a technique that has not been extensively studied. In particular, the relative contributions of moving, real-time object shadows in facilitating performance with interactive, spatial tasks has not been fully examined. In part, this is because contemporary hardware has just recently become powerful enough to support practical applications involving the computation and rendering of real time shadows for user-controlled, moving objects.

The purpose of this study is to investigate how the presence of shadows created by different light sources and cast by objects on different backgrounds provides depth information about the relative placement and size of objects in 3D space. It is of particular interest to compare the relative contributions of object shadows with stereoscopic viewing in benefiting interactive task performances involving the positioning and resizing of objects in 3D space. This article reviews the more powerful 3D depth cue display techniques and then presents a formal experimental study that examines the power of cast shadows in providing depth information needed to complete 3D object positioning and resizing tasks. The findings are discussed with respect to theory, and future research.

2. THEORY AND BACKGROUND

2.1 Theory

Theoretical studies on the mental representations of visual graphic images have examined the perception of multidimensional stimuli. A *multidimensional* computer-generated image is one that combines several realism (or depth) cues (Lockhead [1972] called them ‘stimulus dimensions’), such as light sources, shadows, surface shading, texture, color, and so forth. An important question is whether these stimulus dimensions are perceived as separate elements, or whether they are perceived as belonging to an integrated, unitary whole object [Garner and Felfoldy 1970; Foard and Nelson 1984].

Lockhead [1972] suggested that *integral stimuli* consist of multiple stimulus dimensions on which the observer can operate simultaneously. Shepard [1964] noted that integral stimuli are those that are related as homogeneous, unitary wholes. An example is color which varies in hue, brightness and saturation. Shepard [1964] described *non-integral stimuli* as those that are intuitively perceived and analyzed as distinct properties, such as size and brightness. For integral stimuli, the combined depth cues will foster the perception of a coherent ‘whole’ image, whereas, for non-integral stimuli, they will not. Accordingly, research on the mental representations of computer-based graphics should address the issue of which combinations of depth cues foster the perception of integral or non-integral images.

Cue theory suggests that the visual system implicitly computes the distances of environmental objects on the basis of information about the posture of the eyes and about patterns of light projected onto the retinas. While it is important to understand how individual depth cues contribute to the holistic perception and interpretation of depth information in computer-rendered scenes, it is equally important to understand how the visual system integrates this information to foster a singular, stable perception of three-dimensionality. It is also relevant to consider whether the depth information provided by the different cues is complementary or conflicting in nature. Several models have been proposed, including the *additive* and *multiplicative* models; the *vetoing hypothesis*; and *weak fusion* and *strong fusion*.

Bruno and Cutting [1988] conducted experiments in which depth information was provided by several non-conflicting pictorial cues, including relative familiar size, height, occlusion and motion parallax. Although these cues provided redundant depth information, subjects always perceived more depth in displays characterized by the presence of more cues. They concluded that depth information from various cues is combined in an *additive* fashion. That is, the combined effects of multiple depth cues can be determined as the simple summation of the information provided by each individual cue. In contrast are *multiplicative* models which suggest that combined depth cue effects are synergistic and greater than the simple summation of each individual cue effect. Sollenberger [1993] reported evidence for the multiplicative model in experiments combining motion parallax and binocular disparity depth cues.

Johnston et al. [1993] suggested a *vetoing* mechanism in situations where strongly conflicting information is provided by different cues. They suggested that in these situations, the stronger, or more dominant depth cue, simply overrides the effect of the weaker cue such that the combined perception of depth is equivalent to that perception of depth resulting from the stronger cue alone. Johnston et al. [1993] cited experiments by Bülthoff and Mallot [1988] in which the depth information provided by the zero retinal disparity cue vetoed depth information provided by shading, resulting in the perception of a flat surface.

Johnston et al. [1993] attributed the *weak fusion* and *strong fusion* mechanisms for cue combination to Clark and Yuille [1990]. The *weak fusion* model suggests that depth information is processed separately by each cue, and then combined in a weighted linear fashion to compute the overall effect. Young et al. [1993] proposed a *modified weak fusion* model in which independent depth estimates formed on the basis of separate depth cues are weighted, and then additively combined by a factor reflecting their “apparent reliability” in that particular visual context. In contrast to *weak fusion* models, the *strong fusion* model suggests that depth cues interact non-linearly. For example, one cue may work to disambiguate information, enabling depth information to be derived from another cue.

Two cues that have been demonstrated to be very powerful in providing depth information are binocular disparity and motion. *Binocular disparity* is a primary physiological cue that enables the perception of depth through stereoscopic images (i.e. stereopsis). Human eyes are located approximately six centimeters apart so that an object viewed with both eyes within a distance of about thirty meters projects two disparate images onto each retina. The brain fuses these two images, and in the process, is provided with significant spatial information about the relative size, shape, location and orientation of the object.

2.2 Dominant Depth Cues

Stereopsis is a powerful depth cue, particularly for objects that are relatively close to the viewer [Yeh 1993]. However, we perceive depth, and the relative positions of objects in space for even the most distant objects. Thus, stereopsis is not the only mechanism for perceiving depth. Numerous studies have investigated human performance capabilities with stereoscopic user interfaces in various task domains, including: cockpit situational awareness [Bridges and Reising 1987; Andre et al. 1990; Reising and Mazur 1990; Yeh and Silverstein 1990]; the viewing, manipulation, grasping and/or recognition of object images [Gallimore and Brown 1993; Hubona et al. 1997; Wickens et al. 1994; Brown and Gallimore 1995; Ware and Franck 1996; Zhai et al. 1996]; relative depth perception [Reinhart 1990]; and medicine [Sollenberger and Milgram 1989]. The use of stereoscopic viewing did not uniformly assist performance in these studies, although several report that stereopsis especially helps when augmenting other monoscopic visual depth cues.

Studies have shown stereo viewing to be a powerful technique for displaying depth information [Hubona et al. 1997; McAllister 1993; Wickens et al. 1989; Yeh and Silverstein 1990] and stereo viewing has been used as a control condition, or baseline, for studying other 3D display techniques [Sollenberger and Milgram 1993; Arthur et al. 1993]. Many of the studies do indicate the benefits of stereoscopic viewing in perceiving, recognizing, grasping and/or understanding object shapes [Hubona et al. 1997; McWhorter et al. 1991; Sollenberger and Milgram 1993; Brown and Gallimore 1995; Ware and Franck 1996; Zhai et al. 1996], although some of the studies do not support the superiority of stereopsis [Gallimore and Brown 1993] as a depth cue and others suggest that these benefits are task specific [Liu and Wickens 1992; Wickens et al. 1994]. However, stereopsis does provide depth cues about object shape that are absent without stereopsis. The general consensus is that stereopsis is a powerful (perhaps dominant) cue for providing the viewer information about the relative location, size, shape, and orientation of objects in 3D space.

Another powerful depth cue is *motion*. Everyday perceptual experiences occur within a context of nested motions. Moving eyes, and moving objects, provide powerful perceptual cues about the environmental and spatial properties of perceived objects. Sollenberger and Milgram [1993] demonstrated the utility of *motion parallax* in graphically visualizing complex, simulated blood vessel structures in the brain. Overbeeke and Stratmann [1988] and Smets [1992] demonstrated that the motion cue can be introduced through the observer’s own head movement. Arthur et al. [1993] and Ware and Arthur [1993] found that the use of head-tracking displays (to introduce the motion cue) had effects as powerful as stereo viewing on task completion times. Moreover, they reported that the effects of head-tracking were *more* powerful than stereoscopic viewing on reducing error rates.

Proffitt and Kaiser [1991] discuss *minimal stimulus conditions* necessary to perceive environmental properties in objects. They maintain that motion is a minimally sufficient condition for perceiving a variety of environmental properties, including 3D form. Although the preponderance of evidence indicates that motion facilitates object recognition performance [Wallach and O’Connell 1953; Braunstein 1976; Todd 1985; Gallimore and Brown 1993; Sollenberger and Milgram 1993], there is evidence to suggest that it does not matter whether the object or the observer is moving [Van Damme and Van de Grind 1993]. Further, although there is some evidence that the rate of motion affects the ability to recognize the object [Sollenberger and Milgram 1993], studies have reported conflicting findings whether motion controlled by the observer is superior to uncontrolled motion [Hubona et al. 1997] or whether the type of motion makes no difference with respect to performance [Ware and Franck 1996].

2.3 Shadows as Spatial Cues

Yonas [1979] noted the distinction between two types of object shadows: the primary or *attached shadow* that occurs when the shadow of an object is visible on that same object, and the derived or *cast shadow* that occurs when the shadow of one object is visible on a different object. Attached shadows, sometimes referred to as *object shading*, are produced by the different surface orientations of an object that self-shade other surfaces of the same object relative to some light source. In contrast to *attached shadows*, *cast shadows* are caused by the absence of light created by an occluding object positioned between the light source and a separate, detached, and otherwise illuminated, object surface.

Yonas [1979] cited a number of studies demonstrating that shadows attached to an object can influence the perceived shape of that object. For example, the human visual system ‘assumes’ that unseen light sources come from ‘above,’ which is typically true in terrestrial environments. As a result, the objects depicted in Figure 1 are usually perceived as concave ‘dimples’ and convex ‘bumps’ on a surface. That is, the top left object in Figure 1 is seen as a depression receding into the surface whereas the middle top object is seen as a round extrusion above the surface. Although the direction of illumination is ambiguous and could originate from either the top or the bottom of the picture, the perceived shape of the object is based on the unconscious assumption that the light is coming from above. When the Figure is rotated 180 degrees, the objects ‘reverse’ in shape such that those seen as concave are now seen as convex and vice versa.

Fig. 1. Attached shadows influence perceived object shape.

It has also been demonstrated that shadows cast by an object can influence the perceived size, elevation and relative depth of that object. Yonas et al. [1978] explored the effects of varying ground plane locations of shadows cast by objects. In Figure 2, most people ‘see’ the right object as larger in absolute size and the left object as farther off the ground and closer to them. Yonas’ studies demonstrate that attached and cast shadows constitute important sources of information about the shapes, sizes, and layouts of objects in space.

Fig. 2. Cast shadows influence perceived object size, elevation and depth.

Wanger et al. [1992] manipulated six depth cues while subjects performed object positioning, rotating, and resizing tasks. The manipulated depth cues included: projection (orthographic or perspective); objects casting shadows (on/off); object texture (checkerboard/solid); ground texture (checker/solid); viewpoint motion (on/off); and elevation (object on ground plane/floating in air). Relative to the other five depth cues, the use of shadows had a dominant effect on enhancing the accuracy of performing the object positioning and resizing tasks.

Although Wanger et al. [1992] demonstrated that positioning and sizing objects in 3D space is aided by the presence of shadows cast by the objects, many applications cannot use computer-generated shadows because they are computationally very expensive to render in real-time. Accordingly, the effects of *shadow quality* on the perception of objects in space becomes an important issue. Wanger [1992] conducted three experiments to investigate the effects of *shadow* *sharpness* and *shadow shape* on the accuracy of spatial estimation tasks. In the first experiment, Wanger investigated the effects of shadow sharpness and shape on perceived object size and position in a fixed scaling task. Although the presence of the shadows aided performance in accurately estimating object height and depth, shadow sharpness (i.e. the accurate rendering of both umbral and penumbral regions) had no effect. The second experiment tested the effect of shadow shape (i.e. 'boxy' versus 'true' shadows) on the perception of object size and position. Shadow shape had no effect. The third experiment tested the effect of shadow sharpness on perceived object shape in a shape matching task. 'Soft' shadows had a significant effect on *impairing* shape matching accuracy. He concluded that although shadows can be a useful cue to indicate an object's 3D shape, soft shadows can be *detrimental* to determining object shape in the absence of other cues. The implication is that the computationally less expensive 'hard shadowing' computer rendering techniques may be more useful in indicating an object's 3D shape than the computationally more expensive 'soft shadowing' techniques.

In addition to Wanger et al. [1992], several studies have noted that the perceived realism of computer-generated images is not directly related to the increased detail of lighting, shading and/or shadowing techniques. Atherton and Caporeal [1985] varied the number of polygons and the type of shading (e.g. flat, Gouraud, Phong) used to generate images of a sphere. They reported that the perceived realism of the sphere increased significantly up to a point, followed by a range of diminishing returns, such that large increases in CPU time produced little or no increased levels of perceived realism. Meyer and Greenberg [1980, 1985], Meyer et al. [1986], and Immel et al. [1986] studied the Lambertian radiosity method of modeling reflected light intensities (in which all light reflection is assumed to be diffuse). Subjects were unable to discriminate between a simple scene consisting of block-like objects rendered using the radiosity method with a photograph of the same scene. That is, the computer-based image was as realistic to the subjects as the photograph.

Todd and Mingola [1983] examined the effects of several realism and depth cues, including shading, texture, specular highlights, and directions of light sources on the perception of the 3D structure of a cylinder. They reported that cylinder surface 'shininess' enhanced the perception of curvature, but was not related to the perception of the direction of illumination. They also reported that shading was less effective than the gradient produced by a checker-board texture in depicting the 3D surface shape.

Wallach and O'Connell [1953] demonstrated that people can recover 3D form when viewing 2D shadows of moving 3D wire form objects projected onto screens. When the wire forms were stationary, the viewers reported seeing only 2D configurations of lines in the shadow projections. However, when the wire forms were continuously rotated, viewers could accurately recognize the 3D objects represented by the wire forms. Wallach and O'Connell labeled this phenomenon the Kinetic Depth Effect, or KDE. The KDE is well documented as a powerful depth cue and has been demonstrated to exist under relatively impoverished conditions, such as when viewing projections of rotating dot patterns [Braunstein 1976; Todd 1985].

Herndon et al. [1992] demonstrated the capability of shadows to reflect an object's shape by presenting a set of 3D 'shadow widgets' that served as direct manipulation interfaces for the corresponding objects. The objects' shadows could be directly manipulated by users to translate, rotate and scale *the objects* that were casting the shadows. The geometries of the shadow widgets were directly related to the geometries of their respective 3D objects. The shadows were constrained to lie in a single plane. Transforming the shadow widget affected the 3D object in the same plane. Users could translate and scale objects by manipulating the object's shadow in the projection plane. Object rotation could be accomplished by manipulating the shadow about the normal to the projection plane. Herndon stated that these shadow widgets reduced the ‘cognitive distance’ of affecting changes to the corresponding objects because they relied on a concrete visual, real-world metaphor conveying their function.

Barfield et al. [1988] varied the number of light sources (one, two), angle of separation between objects (0^{o}, 60^{o}, 120^{o}, and 180^{o}), object complexity (simple, complex) and the type of surface shading (flat, smooth) in a ‘same-different’ object matching mental rotation task. The number of light sources and the type of surface shading did not affect mean response times. However, wider angles of separation between the target and comparison objects resulted in slower response times. Mean response times for simple objects were faster than for complex objects.

Pani et al. [1996] conducted experiments investigating whether variations in object orientation affect the ability to imagine the structure and shape of that object’s shadow. They had subjects imagine and draw the shadows that would be cast by visible objects that varied in orientation in relation to the direction of projection (i.e. the light source). With basic rectilinear and simple platonic solids, subjects were successful at imagining the shape and structure of shadows when the objects were aligned (i.e. parallel or perpendicular) with the direction of the light source. As the angle became more oblique, performance deteriorated rapidly.

Zhai et al. [1996] described the *partial-occlusion* effect which applies the concept of semitransparent surfaces as an effective depth cue in 3D target localization tasks. *Occlusion*, or the presence of features hidden behind an object (sometimes referred to as *interposition*), has long been recognized as a powerful depth cue [Schriever 1925; Braunstein et al. 1986; Wickens et al. 1989; Gallimore and Brown 1993; Brown and Gallimore 1995]. Hidden features allow the viewer to infer which surface is ‘in front.’ However, occlusion is difficult to use with 3D interaction tasks because distal objects may be completely obscured by the more proximal, opaque objects, thereby creating uncertainties about unseen background objects. Partial occlusion overcomes this inherent limitation of total occlusion and still provides depth information. Partial occlusion, sometimes referred to as the *silk* effect, is a type of shadowing technique. When a semitransparent surface overlaps another object, the overlapped object is seen in lower contrast (i.e. as if through a silk stocking). Depth information is provided by the ‘darkness’ of the partial occlusion: the darker the partial occlusion, the more distant the object.

Fig. 3. Example of positioning task trial.

3. METHOD

3.1 Experimental Tasks

In the current study, subjects performed separate object positioning and object resizing tasks.

3.1.1 *Positioning Task.* Subjects were successively presented with object images, including cubes, spheres, and tetrahedrons, arranged to outline the vertices of symmetrical figures, including cubes and octahedrons. One of the vertex objects was deliberately misplaced in the symmetrical arrangement. The task was to reposition the misplaced vertex object in 3D space (i.e. in the x, y and z dimensions), as quickly and accurately as possible, so as to complete the symmetrical arrangement. Figure 3 depicts a positioning task trial using cube-shaped vertices arranged in an overall cube figure over a plane illuminated by one light source. In this particular trial, the subject's task was to move the cube in the center of the display down to the lower front left corner (or vertex) of the overall cube arrangement. Subjects repositioned the vertex objects using a *spaceball*, a six-degrees-of-freedom input device.

Fig. 4. Example of resizing task trial.

3.1.2 *Resizing Task*. In the resizing task, subjects were again presented with object images (cubes, spheres, or tetrahedrons) arranged to outline the vertices of a symmetrical figure (a cube or an octahedron) in 3D space. Unlike the positioning task, in the resizing task, the vertex objects were correctly positioned to complete the symmetrical arrangement. However, one of the vertex objects was a different size (either smaller or larger) than the remaining objects. The task was to resize this mismatched object, as quickly and accurately as possible, to make it correspond with the uniform size of the other objects. Figure 4 depicts a resizing task trial using spheres with no light sources arranged in an octahedron over stairs. Subjects again used the spaceball input device to resize the objects.

3.2 Experimental Conditions and Hypotheses

There are four experimental conditions in this experiment: *shadows* (on, off); *number of light sources* in the shadows-on condition (one or two); *viewing mode* (stereo, mono); and shadow *background* (flat plane, room, 'stair-step' plane). Of primary interest are the relative contributions of the *shadows on* and *stereo viewing* conditions in assisting task performance. Accordingly, with respect to both tasks, the following hypotheses are proposed:

1. The presence of objects casting shadows improves performance in spatial tasks.
2. Stereoscopic viewing also improves performance in spatial tasks.

Of secondary interest are performance effects arising from the remaining conditions. An important theoretical issue is that of *dimensional integrality*. Specifically, when scientific data contain more than two dimensions, how are they best rendered so as to facilitate comprehension? There have been many suggested techniques: XY plots with Z represented through intensity or color [Liu and Wickens 1992] or with surface contours [Van Damme and Van de Grind 1993]; by the concurrent use of two, or even three, 2D planar representations [Andre et al. 1990; Wickens et al. 1994]; and 3D perspective drawings [Carswell et al. 1991; Liu and Wickens 1992]. Wickens [1992a and 1992b] and Wickens et al. [1994] argue that the *proximity compatibility principle* asserts there is an inherent advantage of an additional display dimension (e.g. a 3D over two planar 2D displays, or an XY plot over two X plots) when multiple sources of data must be integrated.

We argue that the *proximity compatibility principle* also applies to spatial tasks such that performance will improve when: (1) the number of light sources increases from one to two; and (2) the background changes from a flat plane to a 'stair-step' plane to a room with walls. Additional light sources cast additional shadows for any given object. Each shadow should provide additional depth information about the corresponding object. Thus, additional depth information should be provided by the extra light source. Similarly, multidimensional background planes also provide additional display dimensions, and consequently, additional depth information. Thus, spatial task performance should improve.

3.3 Experimental Design and Procedure

The experiment used a 2 x 2 x 2 x 3 within-subjects design, manipulating the independent variables: *shadows* (on, off); *number of light sources* (one, two); *viewing mode* (stereo, mono); and shadow *background* (flat plane, room, 'stair-step' plane). There were a total of four sets of 72 trials each, or 288 trials per subject, 144 trials for each task. All subjects viewed the same 288 scenes, although the presentation order varied.

One half of all trials randomly presented shadows on, and one half, shadows off. In the shadows on condition, one half of the trials randomly used one light source, and the other half used two light sources. One half of the trials were viewed in stereo, and the other half in mono. The stereo and mono conditions were viewed in 'blocks' of 72 trials each because the subjects wore special eyeglasses to produce the stereo effect. One third of the background conditions randomly used a flat plane for the shadow background; one third used the room; and the remaining third used the 'stair-step' plane. Furthermore, the orientation of the stair-step plane was randomly varied, left-to-right, front-to-back, and right-to-left. Figure 3 presents an example of the flat plane shadow background. Examples of the (front-to-back) 'stair-step' plane, and room shadow backgrounds are respectively presented as Figures 5 and 6.

All light sources were positioned 'above' the vertex objects so that their shadows were visible on the background surface. Moreover, the light sources were placed at various angles from the zenith. In the case of two light sources, their angle of separation varied from trial to trial. Thus, the relative locations and separation of the light sources changed from trial to trial. In both the positioning and resizing task trials, the location of the misplaced or missized vertex was randomly placed with respect to the remaining vertices.

Fig. 5. ‘Stair-step’ shadow background.

The dependent variables included *error magnitude* and *response time*. Repeated measures of the dependent variables were automatically recorded by the task trial software. Each subject participated in a limited number of practice trials to familiarize them with the positioning and resizing tasks, and with the use of the spaceball input device. Following the practice trials, each subject engaged in 2 sets of 72 trials for each of the positioning and resizing tasks, for a total of 144 observations per task per subject. One half of the subjects completed the positioning trials first, and one half completed the resizing task trials first. Subjects were permitted to rest briefly between each set. Subjects completed all positioning and resizing task trial sets in one session that ranged in length from 75 to 120 minutes.

The task trial software was developed on a Silicon Graphics, Inc. (SGI) Onyx2 workstation with Infinite Reality2 hardware and IRIX 6.3 system software. The code was written in C and uses OpenGL and GLUT 3.5. Based on the GLUT demo shadowmap.c file provided by Tom McReynolds of SGI, the code implements cast shadows using projective textures. The scene is rendered from the viewpoint of each light position. The resulting depth map is saved into the texture image. The scene is again rendered from the user's viewpoint, and the saved texture becomes the cast shadows.

Fig. 6. Room shadow background.

Stereo was implemented using the older method common on SGIs: two buffers represent the top and bottom halves of the display area. The scene is rendered into the top back buffer using a stereo projection algorithm included with the GLUT demos, then rendered into the bottom back buffer. Then the front and back buffers are swapped. Stereoscopic viewing was enabled by wearing 120Hz flicker-free stereoscopic CrystalEyes™ glasses (model no. CE-1), manufactured by StereoGraphics.

As each scene was presented, the target vertex object for positioning or resizing was identified by a reddish color whereas the remaining objects were blue. As soon as the subject manipulated the object with the spaceball, the reddish color changed to match the blue color of the other objects. When finished positioning or resizing the target object, the subject pressed a button on the spaceball pad, which caused the results of the trial to be recorded and the next randomly selected scene in the trial set to be presented.

The subject population consisted of thirty employees and contractors of the Goddard Space Flight Center. Subjects voluntarily participated in the experiment. All subjects had professional occupations as engineers, computer programmers or computer scientists. Each subject completed a preliminary questionnaire soliciting demographic information. The mean age of subjects was 40.2 years, with 5.6 mean years of education beyond high school, 15.42 mean years of computer experience, and 18.5 mean years of professional work experience.

2.4 Performance Measures

Two task performance measures were collected: *error magnitudes* and trial completion times (or *response times*). The base unit of measurement for error magnitude was the standard measurement unit in x, y and z space provided by the software. Error magnitude for the positioning task was defined as the Euclidean summation of the three directional errors in the x, y, and z dimensions (i.e. error magnitude = $(e^2_x + e^2_y + e^2_z)^{1/2}$). This metric constitutes the exact distance of the repositioned object from its correct position in three dimensional space. Error magnitude for the scaling task was defined as the absolute value of the error in length of either the radius (in task trials using spheres as vertex objects) or the diagonal (in task trials using cubes or tetrahedrons as vertex objects) of the resized object.Trial completion time was the period of time, measured in milliseconds, from when the scene first appeared until the subject pushed a button on the spaceball, causing the next scene to appear. They would push the spaceball button to indicate that they had completed that particular trial.

2.5 Experimental Results

Table I displays the means and standard deviations of error magnitudes and response times for the two tasks. For each task, the performance measures are displayed by use of shadows (on, off), viewing mode (stereo, mono), number of light sources (one, two), and background (flat plane, ‘stair-step’ plane, room). For the positioning task, the mean overall error magnitude was 1.316 and the mean overall response time was 16.108 seconds. For the resizing task, the mean overall error magnitude was 0.080 and the mean overall response time was 8.985 seconds.

Table I. Means and standard deviations of error magnitudes and response times.

	Positioning Task				Scaling Task			
	Error Magnitude		Response Time (secs.)		Error Magnitude		Response Time (secs.	
Use of Shadows:	Mean	Std dev	Mean	Std dev	Mean	Std dev	Mean	Std d
On	1.225	1.399	16.634	12.025	0.079	0.099	9.198	5.74
Off	1.406	1.618	15.581	11.202	0.081	0.096	8.773	5.56
Viewing Mode:	Mean	Std dev	Mean	Std dev	Mean	Std dev	Mean	Std d
Stereo	0.737	0.632	14.125	9.838	0.073	0.090	8.559	5.63
Mono	1.894	1.877	18.091	12.883	0.088	0.103	9.411	5.64
Number of Lights:	Mean	Std dev	Mean	Std dev	Mean	Std dev	Mean	Std d
One	1.129	1.291	15.623	10.490	0.076	0.085	8.924	5.37
Two	1.316	1.489	17.590	13.249	0.082	0.110	9.458	6.07
Shadow Background:	Mean	Std dev	Mean	Std dev	Mean	Std dev	Mean	Std d
Plane	1.279	1.408	15.783	10.913	0.087	0.092	8.817	5.55
Room	1.225	1.410	15.660	10.803	0.068	0.083	9.048	5.76
Stairs	1.443	1.700	16.880	13.014	0.086	0.113	9.091	5.65

The data was analyzed by fitting a repeated measures multivariate analysis of variance model (MANOVA) to the experimental observations. The MANOVA model tested each of the four main effects (shadows on/off, viewing mode, number of lights, and background) on the error magnitude and response time dependent variables. The data from the two tasks were segregated and analyzed separately. Type III sums of squares are reported for main effects that were also components of significant interaction effects. Type III sums of squares take into account the influence of main effects before the influence of interaction effects are computed. Thus, a significant interaction effect does not nullify the statistical significance of any corresponding main effect.

3.5.1 *Positioning Task Results*. In the positioning task performance data, there were significant differences (at the 95% confidence level) in the mean values of the dependent variables (error magnitude and response time) as a function of all four main effects: use of shadows; viewing mode; number of lights; and shadow background. In addition to these significant main effects, there was also a significant interaction between the number of light sources and the shadow background.

The use of shadows significantly affected subjects' object positioning performances in the omnibus MANOVA model (F (2, 4280) = 19.79; p = 0.0001). With shadows on, subjects were more accurate (the 'shadows on' condition mean error magnitude is 1.225; F (1, 4281) = 19.74; p = 0.0001) than they were without the shadows (the 'shadows off' condition mean error magnitude is 1.406). However, subjects *took longer* positioning the vertex objects using shadows (the mean 'shadows on' positioning response time is 16.634 seconds; F (1, 4281) = 13.59; p = 0.0002) than did subjects without the shadows (the mean 'shadows off' condition response time is 15.581 seconds).

The viewing mode also significantly impacted subjects' positioning task performances in the omnibus MANOVA model (F (2, 4280) = 442.28; p = 0.0001). Subjects viewing the vertex objects in stereo were more accurate (the mean stereo viewing condition error magnitude is 0.737; F (1, 4281) = 804.56; p = 0.0001) than were subjects viewing the vertex objects in mono (the mean mono viewing condition error magnitude is 1.894). Furthermore, subjects viewing the objects in stereo were faster performing the positioning task (mean stereo viewing condition response time is 14.125 seconds; F (1, 4281) = 193.04; p = 0.0001) than were subjects viewing the objects in mono (the mean mono condition response time is 18.091 seconds).

The number of lights also significantly impacted subjects' positioning task performances in the omnibus MANOVA model (F (2, 4280) = 11.50; p = 0.0001). Subjects viewing the vertex objects with two lights sources were *less* accurate (the mean two lights condition error magnitude is 1.316; F (1, 4281) = 6.24; p = 0.0125) than were subjects viewing the vertex objects with one light source (the mean one light condition error magnitude is 1.129). Furthermore, subjects viewing the objects with two light sources took *more* time performing the positioning task (mean two light condition response time is 17.59 seconds; F (1, 4281) = 19.73; p = 0.0001) than did subjects viewing the objects under one light source (the mean one light condition response time is 15.623 seconds).

The shadow background also significantly impacted subjects' positioning task performances in the omnibus MANOVA model (F (4, 8560) = 8.99; p = 0.0001). Subjects viewing the vertex objects in a room (the mean room background condition error magnitude is 1.225; F (2, 4281) = 12.41; p = 0.0001) or over a plane (the mean plane background condition error magnitude is 1.279) were more accurate than were subjects viewing the vertex objects over stairs (the mean stairs condition error magnitude is 1.443). The error magnitude difference between the room and plane conditions was not significant. Furthermore, subjects viewing the objects over stairs (the mean stairs background condition response time is 16.88 seconds; F (2, 4281) = 8.44; p = 0.0002) took more time performing the positioning task than did subjects viewing the objects in a room (the mean room background condition response time is 15.66 seconds) or over a plane (the mean plane condition response time is 15.783 seconds). The difference in the response times between the room and plane background conditions was not significant.

In the positioning task data, there was also a statistically significant interaction between the number of lights and the shadow background in the omnibus MANOVA model ($F(8, 8560) = 2.78$; $p = 0.0045$). This interaction effect was significant with respect to error magnitude ($F(4, 4281) = 4.51$; $p = 0.0012$), but not with respect to response time ($F(4, 4281) = 1.55$; $p = 0.1847$). Figure 7 graphically presents this interaction effect.

Fig. 7. Interaction of number light sources with shadow background on positioning accuracy.

3.5.2 *Resizing Task Results*. In the resizing task performance data, there were significant differences (at the 95% confidence level) in the mean values of the dependent variables (error magnitude and response time) as a function of all four main effects: use of shadows; viewing mode; number of lights; and shadow background. In addition to these significant main effects, there was also a significant interaction between the viewing mode and the number of light sources.

The use of shadows significantly affected subjects' object resizing performances in the omnibus MANOVA model ($F(2, 4282) = 6.61$; $p = 0.0014$). There was no difference in accuracy (the 'shadows on' condition mean error magnitude is 0.079; $F(1, 4283) = 0.71$; $p = 0.3982$; whereas the 'shadows off' condition mean error magnitude is 0.081). However, subjects *took longer* resizing the vertex objects using shadows (the mean 'shadows on' resizing response time is 9.918 seconds; $F(1, 4283) = 12.08$; $p = 0.0005$) than did subjects without the shadows (the mean 'shadows off' condition response time is 8.773 seconds).

The viewing mode also significantly impacted subjects' resizing task performances in the omnibus MANOVA model ($F(2, 4282) = 24.52$; $p = 0.0001$). Subjects viewing the vertex objects in stereo were more accurate (the mean stereo viewing condition error magnitude is 0.073; $F(1, 4283) = 18.59$; $p = 0.0001$) than were subjects viewing the vertex objects in mono (the mean mono viewing condition error magnitude is 0.088). Furthermore, subjects viewing the objects in stereo were faster performing the resizing task (mean stereo viewing condition response time is 8.559 seconds; $F(1, 4283) = 33.50$; $p = 0.0001$) than were subjects viewing the objects in mono (the mean mono condition response time is 9.411 seconds).

The number of lights also significantly impacted subjects' resizing task performances in the omnibus MANOVA model ($F(2, 4282) = 5.17$; $p = 0.0057$). There was no difference in accuracy (the mean two lights condition error magnitude is 0.082; $F(1, 4283) = 1.79$; $p = 0.1805$) than were subjects viewing the vertex objects with one light source (the mean one light condition error magnitude is 0.076). However, subjects viewing the objects with two light sources took more time performing the resizing task (mean two light condition response time is 9.458 seconds; $F(1, 4283) = 9.03$; $p = 0.0027$) than did subjects viewing the objects under one light source (the mean one light condition response time is 8.924 seconds).

The shadow background also significantly impacted subjects' resizing task performances in the omnibus MANOVA model ($F(4, 8564) = 9.96$; $p = 0.0001$), but only with respect to resizing accuracy, not response time. Subjects viewing the vertex objects in a room (the mean room background condition error magnitude is 0.068; $F(2, 4283) = 17.72$; $p = 0.0001$) were more accurate than subjects in either the plane condition (the mean plane background condition error magnitude is 0.087) or the stairs condition (the mean stairs condition error magnitude is 0.086). The differences in the response times among the three shadow background conditions were not significant ($F(2, 4283) = 1.81$; $p = 0.1646$). The mean response time resizing objects over a plane was 8.817 seconds, over a room was 9.048 seconds, and over stairs was 9.091 seconds.

In the resizing task data, there was a significant interaction effect between viewing mode and the number of lights ($F(4, 8564) = 3.201$; $p = 0.0123$) with respect to both accuracy ($F(2, 4283) = 3.22$; $p = 0.0402$) and response time ($F(2, 4283) = 3.58$; $p = 0.0281$). Figures 8 and 9 graphically present these interactions.

4. DISCUSSION

The findings of this study are mixed with respect to the original hypotheses. The presence of object shadows did not uniformly improve performance in both spatial tasks. In the positioning task, accuracy was improved, but response times were longer. In the resizing task, there was no effect on accuracy, and response times were longer.

The presence of the object shadows did improve subjects' accuracy in positioning objects, but at the expense of taking more time to finish the trials. In spite of longer response times with the shadows, one could argue that the *quality* of their positioning task performance was enhanced since they were more accurate. However, they took *longer* to position the objects using the shadows, even though they were instructed to work as accurately *and* quickly as possible. Indeed, in informal conversations with the subjects immediately following the positioning trials, many stated that they only used the additional shadowing cue *after* they had used other perceptual cues to place the mispositioned object in the perceived correct location. That is, they used the shadowing cue as a 'final check' to more precisely calibrate the position of the misplaced vertex object. However, it is relevant to note, that, in these conversations, many subjects expressed great enthusiasm for the utility of the shadows in assisting positioning task performance. Considering the experimental data in conjunction with this anecdotal evidence suggests that many subjects used the shadow cues to more accurately 'fine tune' an initial positioning decision.

However, in the resizing task, the presence of the object shadows did not improve performance accuracy at all, and, they again *took longer* to complete the trials. This degraded performance in resizing the objects in the presence of cast shadows belies the fact that many subjects expressed confidence that the shadows aided their performance. In fact, as a group, they performed better with no shadows at all. Evidently, they again attempted to use the shadowing cue for additional final assistance in calibrating the object's size, but it did not significantly improve their accuracy (even though many thought that it did).

In contrast to the mixed results in subjects' performances tied to the shadowing cue, stereo viewing, compared to mono viewing, uniformly improved all measures of performance for both spatial tasks, as speculated by the hypotheses. Subjects were both faster and more accurate positioning and resizing the objects in stereo. Furthermore, when subjects were asked whether they thought the stereo viewing mode was helpful, they typically responded in the affirmative. Responses such as: "Sure!" and "Of course!" were common. Clearly, the empirical evidence, as well as the subjects' verbal responses, indicate that stereo viewing is a powerful cue that had a dominant influence on the spatial task performances, relative to the shadowing cue.

Why did stereo viewing have such a dominant influence? Kelsey (1993) cites Neisser (1967) in discussing a particular model of the human visual system, that of *preattentive and attentive processing textons*, that might help explain the performance benefits of stereo viewing. He noted that the speed of visual perception implies a *preattentive* level of processing early in the perceptual process, followed by subsequent and more focused *attentive* processing. Preattentive processing is automatic, and very fast, such that visual stimuli are processed in parallel. Certain features called “textons,” such as color, line orientation (i.e. ‘edges’), and line intersections (i.e. ‘corners’), are detected at the preattentive level. Texton differences enable the rapid recognition of a target, to be followed by eye-head movement to bring the target into foveal vision for more focused attentive processing. Attentive processing is more deliberate and requires conscious cognitive comparisons and the use of memory. The introduction of stereo viewing, particularly when viewing objects with edges and corners, makes the distance (i.e. depth) comparison and image sizing tasks easier by shifting more of the perceptual features from the attentive to the preattentive level. Consequently, some of the burden of comparing object positions and sizes is shifted, from the more deliberate and conscious attentive level, requiring the expenditure of mental effort, onto the more automatic and faster preattentive perceptual level. Thus, spatial object positioning and resizing tasks are facilitated.

However, perhaps the data with respect to the number of light sources is more revealing regarding the utility of the shadows in assisting these task performances. The hypotheses speculate that both positioning and resizing task performances would improve as the number of light sources increases from one to two. This was not the case. Positioning accuracy significantly improved as the number of light sources increased from none (mean error magnitude = 1.406) to one (mean error magnitude = 1.129), but then significantly declined from one light source to two (mean error magnitude = 1.316). In fact, positioning accuracy with two light sources was not significantly different from that with no light sources (i.e. no shadows at all). Positioning response times were not significantly different with no lights (mean response time = 15.58 seconds) compared to one light (mean response time = 15.623 seconds), but with two lights (mean response time = 17.59 seconds), subjects were significantly *slower* in positioning the objects than with either one or no lights.

Resizing accuracy was not significantly different with either none (mean error magnitude = 0.081), one (mean error magnitude = 0.076), or two (mean error magnitude = 0.082) light sources. Similarly, resizing response times were not significantly different with no lights (mean response time = 8.773 seconds) or one light (mean response time = 8.924 seconds), although subjects were significantly slower with two light sources (mean response time = 9.458 seconds) than with either one or no lights.

These results indicate that, contrary to our speculation, more light sources did not provide better and more comprehensive spatial information about the objects’ positions and sizes. Our *a priori* assumption that additional light sources would provide linearly increasing amounts of spatial information about the corresponding objects was flawed. In fact, two light sources often fostered *ambiguous* and *conflicting* spatial information about the objects. Recall that the position of the light sources varied from trial to trial. Although subjects reported they could more effectively infer the position of an object with one light source than with none, a number of subjects expressed confusion about the positions of both the lights and the objects when there were two light sources. In these cases, the shadows did not help them infer the relative positions and sizes of the corresponding objects. In ensuing discussions, subjects speculated that two lights might have helped more had they been uniformly fixed in their angular and distance orientations in each scene.

Barfield et al. [1988] reported an interaction with respect to response times between the number of light sources (one or two) and object complexity, but found no difference in reaction times for one or two sources of illumination across each level of complexity. This finding was in spite of their prior prediction that adding light sources would facilitate the forming of propositions and therefore result in reduced reaction times. They stated [p. 681]: “It may also be true that adding two light sources had no effect in making the images more discriminable in some sense, a conjecture supported by the finding that rendering the shaded images with one versus two sources of illumination did not lead to significant differences in response times or response accuracy." However, they were dealing with object self-shading (or attached shadows) rather than with shadows cast by objects onto a detached background.

Similarly, with the shadow background conditions, our results deviated from those hypothesized, perhaps because of the same erroneous assumption of linearly increasing amounts and quality of spatial information as the shadow background becomes more 'multidimensional' (i.e. changes from a plane to stairs to a room). In the positioning task, the stair-step background was problematic for performance, both in terms of accuracy and response time. There was no difference in positioning accuracy, nor response time, for the plane compared to the room. In the resizing task, accuracy was improved with a room background, compared to either the plane or stairs. There were no significant resizing response time differences among either of the three background conditions.

However, recall that the shadows had no beneficial effect in the *resizing* task, although they did help in more accurately *positioning* objects. If the *shadows* have no beneficial effect in resizing objects, then it logically follows that the *shadow background* should also be inconsequential for resizing objects. Thus, we do not attempt to interpret the shadow background results in the resizing task. However, notice the significant interaction (see Figure 7) between the number of light sources and the shadow background in positioning task accuracy. Introducing one light improves positioning accuracy, regardless of shadow background. However, introducing the second light *impairs* performance, greatly with the stairs, but also over a plane. However, over the room, the second light serves to *improve* positioning accuracy.

We argue that positioning the objects in a room does achieve a multidimensional effect consistent with the *proximity compatibility principle* [Wickens 1992a and 1992b; Wickens et al. 1994] such that the task of accurately positioning objects is enhanced. As the objects are moved within the room, there is less ambiguity regarding the 3D positioning of the objects, even in the presence of two light sources. We suggest there is less ambiguity for two reasons. As the objects approach a 'wall' in the room, the shadow associated with that object is immediately apparent as the shadow 'climbs' up or down the wall to 'touch' the object. Thus, the object's relative 3D position is very apparent even if the object has an additional shadow on the floor or elsewhere. Furthermore, as the object itself 'touches' and 'passes into' the wall, the resulting occlusion of part of the object provides an additional powerful depth cue that is not evident by moving objects over a plane or over stairs. Thus, we suggest that bounding the object *inside* a room, with walls in 3D space, provides additional powerful depth cues to locate that object's position relative to other objects in the room.

Fig. 8. Interaction of viewing mode with the number of light sources on resizing accuracy.

5. CONCLUSIONS

This evidence clearly suggests that stereo viewing is more powerful than the use of object shadows in providing depth cues about the relative size and position of objects in space. But how does this evidence relate to alternate versions of *cue theory*? That is, what does this data indicate about how visual depth cues are 'combined' from a human information processing perspective so as to provide an integrated cognitive impression regarding some spatial characteristic of the perceived object, such as relative or absolute size, distance, elevation, and so forth?

Fig. 9. Interaction of viewing mode with the number of light sources on resizing response time.

Figures 8 and 9, which depict the interaction of viewing mode with the number of light sources, provide some clues in this regard. In Figure 8, notice that the introduction of a single light source precipitates improved resizing accuracy in the mono viewing mode, but does not improve accuracy in the stereo mode. Introducing the second light source impairs resizing accuracy in mono, and does not significantly change the level of resizing accuracy in stereo. In Figure 9, a similar pattern emerges with respect to resizing response time. Adding one light, and then two, causes resizing response time to successively increase in the stereo mode. In the mono viewing mode, adding one light reduces the response time, while adding the second light causes response time to increase.

Because this experiment is designed such that viewing mode, and object shadows produced by the light sources, are the dominant depth cues, we can draw some inferences regarding the alternate versions of cue theory. In this experiment, stereo viewing is dominant over the number of shadows. Combining stereo viewing with one, and then two shadows, causes both resizing accuracy and response time to *degrade*. That is, task performance while viewing the objects in stereo *gets worse* as lights are added. This observation clearly does not support either the *additive* [Bruno and Cutting 1988] or *multiplicative* [Sollenberger 1993] versions of cue theory. Each theory predicts improved performance, either additively, or multiplicatively, as more depth cues are added to the scene. This data is more consistent with the *vetoing* and *strong fusion* mechanisms [Johnston et al. 1993]. The *vetoing* mechanism suggests that the more dominant depth cue (i.e. stereo) simply overrides the weaker cue (i.e. shadows), such that the combined effect is no stronger than that produced by the more powerful cue. The *strong fusion* model suggests that depth information is processed separately by each cue, and then combined in a non-linear fashion, that is contingent on the situation and the task. This perspective also helps to explain the differential effects of introducing shadows in the mono and stereo viewing conditions on task performance.

5.1 Future Research

We agree with Wanger et al. [1992] that there is the need for more research to investigate and develop a comprehensive taxonomy of: (1) spatial manipulation tasks; (2) types (i.e. dimensions) of spatial information; and (3) the visual cues that provide this information, such that more effective 3D applications and accompanying user interfaces can be developed. With the demonstrated power of stereo viewing as a dominant cue, more research investigating the efficacy of a wider range of pictorial cues against this baseline cue is warranted. Furthermore, with the advent of the world wide web and the widespread adoption and use of interactive, visual, animation-enabling languages, such as VRML and java, more research is needed to determine how best to display 3D pictures and scenes on the web such that more effective use may be made of this global technology.

ACKNOWLEDGMENTS

This research was sponsored by Code 522.2 of the Goddard Space Flight Center (GSFC) of the National Aeronautics and Space Administration (NASA) in Greenbelt, Maryland, USA.

REFERENCES

ANDRE, A.D., WICKENS, C.D., MOORMAN, L. AND BOSCHELLI, M. 1990. *SID International Symposium Digest of Technical Papers. 21*, 347-350.

ARTHUR, K., BOOTH, K., AND WARE, C. 1993. Evaluating 3D task performance for fish tank virtual worlds. *ACM Transactions on Information Systems. 11*, 3 (July), 239-265.

ATHERTON, P.R. AND CAPOREAL, L.R. 1985. A subjective judgment study of polygon based curved surface imagery. In *Proceedings of SIGCHI, Human Factors in Computing Systems,* 27-34.

BARFIELD, W., SANDFORD, J., AND FOLEY, J. 1988. The mental rotation and perceived realism of computer-generated three-dimensional images. *International Journal of Man-Machine Studies. 29*, 669-684.

BRAUNSTEIN, M.L. 1976. *Depth perception through motion*. New York: Academic.

BRAUNSTEIN, M.L., ANDERSON, G.J., ROUSE, M.W., AND Tittle, J.S. 1986. Recovering viewer-centered depth from disparity, occlusion and velocity gradients. *Perception Psychophys. 40,* 216-224.

BRIDGES, A.L. AND REISING, J.M. 1987. Three-dimensional stereographic pictorial visual interfaces and display systems in flight simulation. *SPIE Proceedings*. *761***,** 102-109.

BROWN, M.E. AND GALLIMORE, J.J. 1995. Visualization of three-dimensional structure during computer-aided design. *International Journal of Human-Computer Interaction*. *7*, 37-56.

BRUNO, N. AND CUTTING, J.E. 1988. Minimodularity and the perception of layout. *Journal of Experimental Psychology: General. 117*, 161-170.

BÜLTHOFF, H.H. AND MALLOT, H.A. 1988. Integration of depth modules: Stereo and shading. *Journal of the Optical Society of America, A. 5,* 10, 1749-1758.

CARSWELL, C.M., FRANKENBERGER, S. AND BERNHARD, D. 1991. Graphing in depth: Perspectives on the use of three-dimensional graphs to represent lower-dimensional data. *Behaviour & Information Technology*. *10*, 459-474.

CLARK, J.J. AND YUILLE, A.L. 1990. *Data Fusion for Sensory Information Processing Systems*. Kluwer Academic Publishers, London..

FOARD, C.F. AND NELSON, D.C.K. 1984. Holistic and analytic modes of processing: the multiple determinants of perceptual analysis. *Journal of Experimental Psychology: General*, *113,* 94-111.

GALLIMORE, J.J. AND BROWN, M.E. 1993. Visualization of 3-D computer-aided design objects. *International Journal of Human-Computer Interaction*. *5*, 361-382.

GARNER, W.R. AND FELFOLDY, G.L. 1970. Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology, 1*, 225-241.

HERNDON, K.P., ZELENIK, R.C., ROBBINS, D.C., CONNER, D.B., SNIBBE, S.S., AND VAN DAM, A. 1992. Interactive shadows. In *Proceedings of the ACM Symposium on User Interface Software and Technology*. ACM, New York 1-6.

HUBONA, G.S., SHIRAH, G.W., AND FOUT, D.G. (in press). The effects of motion and stereopsis on 3D visualization. *International Journal of Human-Computer Studies*.

IMMEL, D., COHEN, M., AND GREENBERG, A. 1986. A radiosity method for non-diffuse environments. In *Proceedings of the 1986 SIGGRAPH Conference, Computer Graphics*, *20*, 133-142.

JOHNSTON, E.B., CUMMING, B.G., AND PARKER, A.J. 1993. Integration of depth modules: Stereopsis and texture. *Vision Research. 33*, 5/6, 813-826.

KELSEY, C.A. 1993. Detection of visual information. In *The Perception of Visual Information*, W.R. HENDEE AND P. WELLS, Eds. New York: Springer-Verlag, 30-51.

LIU, Y. AND WICKENS, C.D. 1992. Use of computer graphics and cluster analysis in aiding relational judgment. *Human Factors*. *34*, 165-178.

LOCKHEAD, G.R. 1972. Processing dimensional stimuli: a note. *Psychological Review*, *79,* 410-419.

MCALLISTER, D.F., Ed. 1993. *Stereo Computer Graphics and Other True 3D Technologies*. Princeton University Press, Princeton, N.J.

MCWHORTER, S.W., HODGES, L.F., AND RODRIGUEZ, W.E. 1991. *Comparison of 3D display formats for CAD applications* (Tech. Report GIT-GVU-91-04). Atlanta: Georgia Institute of Technology, Graphics, Visualization and Usability Center.

MEYER, G.W. AND GREENBERG, D.P. 1980. Perceptual color spaces for computer graphics. *Computer Graphics*, *14*, 254-261.

MEYER, G.W. AND GREENBERG, D.P. 1985. Color education and color synthesis in computer graphics. *Program of Computer Graphics, Cornell University*.

MEYER, G.W., RUSHMEIER, H.E., COHEN, M.F., GREENBERG, D.P., AND TORRANCE, K.E. 1986. An experimental evaluation of computer graphics imagery. *ACM Transactions on Graphics*, *5,* 1, 30-50.

NEISSER, U. 1967. *Cognitive Psychology*. New York: Appleton-Century-Croft.

OVERBEEKE, C.J. AND STRATMAN, M.H. 1988. Space through movement. Ph.D. thesis, Delft Univ. of Technology, Delft, Netherlands.

PANI, J., JEFFRES, J., SHIPPEY, G., AND SCHWARTZ, K. 1996. Imagining projective transformations: Aligned orientations in spatial organization. *Cognitive Psychology*. *31*, 125-167.

PROFFITT, D.R. AND KAISER, M.K. 1991. Perceiving environmental properties from motion information: Minimal conditions. In S. ELLIS, Ed. *Pictorial Communication in Virtual and Real Environments*. Bristol, PA: Francis & Taylor, Inc., 47-60.

REINHART, W.F. 1990. *Effects of depth cues in depth judgments using a field-sequential stereoscopic CRT display*. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

REISING, J.M. AND MAZUR, K.M. 1990. 3-D displays for cockpit: Where they payoff. *SPIE Proceedings*. *1256*, 35-43.

SCHRIEVER, W. 1925. Experimentelle studien über das stereoskopische sehen. *Zeitschrift für Pscyhologie 96*, 113-170.

SHEPARD, R.N. 1964. Attention and the metric structure of the stimulus space. *Journal of the Mathematical Psychology*, *1*, 54-87.

SMETS, G.J.F. 1992. Designing for telepresence: The interdependence of movement and visual perception implemented. In *Proceedings of 5th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Man-Machine Systems*, Pergamon Press.

SOLLENBERGER, R.L. 1993. *Combining depth information: Theory and implications for design of 3D displays*. Ph.D. thesis, Dept. of Psychology, Univ. of Toronto, Ontario, Canada.

SOLLENBERGER, R.L. AND MILGRAM, P. 1993. Effects of stereoscopic and rotational displays in a three-dimensional path-tracing task. *Human Factors*. *35*, 483-499.

TODD, J.T. 1985. Perception of structure from motion: Is projective correspondence of moving elements a necessary condition? *Journal of Experimental Psychology: Human Perception & Performance*. *11*, 689-710.

TODD, J.T. AND MINGOLA, E. 1983. Perception of surface curvature and direction of illuminations from patterns of shading. *Journal of Experimental Psychology*, *9*, 583-595.

VAN DAMME, W.J.M. AND VAN DE GRIND, W.A. 1993. Active vision and the identification of three-dimensional shape. *Vision Research*. *11*, 1581-1587.

WALLACH, H. AND O'CONNELL, D.H. 1953. The kinetic depth effect. *Journal of Experimental Psychology*. *45***,** 205-217.

WANGER, L. 1992. The effect of shadow quality on the perception of spatial relationships in computer generated imagery. In *Proceedings of the 1992 ACM SIGGRAPH Symposium on Interactive 3D Graphics*. *25,* 2 (March), ACM, New York, 39-42.

WANGER, L. FERWANDA, J., AND GREENBERG, D. 1992. Perceiving spatial relationships in computer-generated images. *IEEE Computer Graphics & Applications*. (May), 44-58.

WARE, C. AND ARTHUR, K. 1993. Fish tank virtual reality. In *Proceedings of INTERCHI'93: ACM Conference on Human Factors in Computing Systems.* ACM, New York, 37-42.

WARE, C. AND FRANCK, G. 1996. Evaluating stereo and motion cues for visualizing information nets in three dimensions. *ACM Transactions on Graphics*. *15,* 121-140.

WICKENS, C.D. 1992a. *The proximity compatibility principle: Its psychological foundation and its relevance to display design*. Tech. Report ARL-92-5/NASA-92-3. Savoy, IL: University of Illinois Institute of Aviation, Aviation Research Lab.

WICKENS, C.D. 1992b. Virtual reality and education. In *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, 842-847.

WICKENS, C.D., MERWIN, D.H. AND LIN, E.L. 1994. Implications of graphics enhancements for the visualization of scientific data: dimensional integrality, stereopsis, motion, and mesh. *Human Factors*. *36*, 44-61.

WICKENS, C.D., TODD, S., AND SEIDLER, K. 1989. Three-dimensional displays: Perception, implementations and applications. CSERIAC Tech. Rep 89-001, Wright Patterson Air Force Base, Ohio.

YEH, Y. 1993. Visual and perceptual issues is stereoscopic display. In *Stereo Computer Graphics,* D.F. McAllister, Eds. Princeton University press, Princeton, N.J. 50-70.

YEH, Y. AND SILVERSTEIN, L.D. 1990. Visual performance with monoscopic and stereoscopic presentations of identical three-dimensional visual tasks. *SID International Symposium Digest of Technical Papers. 21,* 359-362.

YONAS, A. 1979. Attached and cast shadows. In *Perception and Pictorial Representation*, C.F. Nodine, Ed. Praeger, New York, 100-109.

YONAS, A., GOLDSMITH, L.T., AND HALLSTROM, J.L. 1978. Development of sensitivity to information provided by cast shadows in pictures. *Perception. 7,* 333-341.

YOUNG, M.J. LANDY, M.S., AND MALONEY, L.T. 1993. A perturbation analysis of depth perception from combinations of texture and motion cues. *Vision Research. 33*, 18, 2685-2696.

ZHAI, S., BUXTON, W., AND MILGRAM, P. 1996. The partial-occlusion effect: Utilizing semitransparency in 3D human-computer interaction. *ACM Transactions on Computer-Human Interaction. 3*, 3 (Sept), 254-284.